

Der Einsatz von PC-Computergraphik in den Sozialwissenschaften: Hardwarevoraussetzungen - Programmkonzepte - Programmbeispiele

Ritter, Heiner

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Ritter, H. (1989). Der Einsatz von PC-Computergraphik in den Sozialwissenschaften: Hardwarevoraussetzungen - Programmkonzepte - Programmbeispiele. ZUMA Nachrichten, 13(24), 42-59. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-209966>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. By using this particular document, you accept the above-stated conditions of use.

Der Einsatz von PC-Computergraphik in den Sozialwissenschaften

Hardwarevoraussetzungen - Programmkonzepte - Programmbeispiele

Von Heiner Ritter

Mit der rasanten Entwicklung der PC-Hardware und -Software stehen mittlerweile auch im Graphikbereich zusätzliche Hilfsmittel für die sozialwissenschaftliche und statistische Analyse zur Verfügung. Die Fülle der angebotenen Produkte und die unterschiedlichen Konzepte machen den Markt jedoch unübersichtlich. Im vorliegenden Beitrag wird anhand einer Checkliste erläutert, was im Zusammenspiel von Graphiksoftware und -hardware zu beachten ist. Am Beispiel von fünf Produkten (SAS/GRAPH, SIR/GRAPH, SPSS/PC+ Graphics mit Microsoft CHART, STATA/GRAPHICS und STATGRAPHICS) werden die unterschiedlichen Konzeptionen der Einbindung von Graphik in die statistische Analyse beschrieben und das Angebot dieser Produkte vorgestellt.

1. Einführung

Mit der Entwicklung der PC-Hardware, den daran angeschlossenen Peripheriegeräten und der PC-Software stehen heute Mittel für die statistische Analyse zur Verfügung, die vormals meist nur einem kleinen Benutzer- und Expertenkreis über Mainframes zugänglich waren. Was zuvor nur mit einem Plotter im Operatorraum eines zentralen Rechenzentrums zu Papier gebracht werden konnte, kann heute mit leistungsfähigen PCs direkt am Bildschirm sichtbar gemacht und mit Tischplottern, hochauflösenden Matrix- und Laserdruckern oder sonstigen Bildaufnahmegeräten meist direkt am Arbeitsplatz weiterverarbeitet werden. Die Spannweite der Einsatzmöglichkeiten reicht vom ersten Entwurf, als Hilfsmittel für weitere Analyseüberlegungen, bis zur fertigen Druckvorlage über Desk-Top-Publishing.¹⁾

Auch im Bereich der Software finden wir mittlerweile ein fast unübersehbares und variantenreiches Angebot. So bieten die klassischen Programmsysteme, wie SAS oder SPSS, auch bei ihren PC-Programmen spezielle Graphikkomponenten an, entweder als integrierte Programmbestandteile oder als Schnittstellen. Darüber hinaus wurden spezielle Produkte entwickelt, die ihren Ausgangspunkt in der graphischen Präsentation und Analyse haben, aber auch Kapazitäten im Bereich des Datenmanagements bieten. Weitere Möglichkeiten ergeben sich durch den Einsatz von kartographischen Systemen, als zusätzliche Dimension der Datenpräsentation und bis in den Bereich der exploratorischen Datenanalyse hinein, die zudem Möglichkeiten für die dynamische Drehung von Punktwolken am Bildschirm bieten.

Mit dieser beeindruckenden Vielfalt beginnt jedoch die Qual der Wahl. Dies betrifft sowohl Kriterien wie die Leistungsfähigkeit, den Preis, die Erlernbarkeit und Bedienungsfreundlichkeit als auch technische Kriterien, wie z.B. die Kernspeicherkapazität, die Menge der Graphikkarten oder die Auflösungsmöglichkeit des Bildschirms. Auch das Verhältnis von (vorhandener) Hardware und Software spielt eine wichtige Rolle in der Entscheidung für ein bestimmtes Programm.

Der vorliegende Artikel gibt eine kurze Zusammenstellung bisheriger Erfahrungen für potentielle Anwendungen sozialwissenschaftlicher Graphikprogramme. Anhand von Beispielen werden Kriterien für die Beurteilung von Graphiksystemen für die sozialwissenschaftliche und statistische Anwendung aufgezeigt. Der Bogen reicht vom Zusammenwirken von Software und Hardware bis zur Beschreibung unterschiedlicher Graphikprogrammkonzepte. Auf der Ebene der Hardware haben wir uns auf IBM- und -Kompatible PCs mit dem DOS-Betriebssystem beschränkt, bei der Software konzentrieren wir uns auf die Graphikprogramme bzw. Graphikprogramm-Module SAS/GRAPH, SIR/GRAPH, SPSS/PC+ Graphics mit CHART, STATA sowie STATGRAPHICS.

2. Wie kommt eine Graphik auf den Bildschirm?

Um eine Graphik auf einem Bildschirm sichtbar zu machen, muß ein entsprechendes Graphikprogramm mit einer PC-Hardware zusammenwirken. Damit sind zwei unterschiedliche Ausgangsbedingungen möglich: Entweder die Software ist vorhanden und die Hardware fehlt oder die Hardware ist vorhanden und die entsprechende Software muß angeschafft werden. Die Lösung solcher Probleme, d.h. die Anschaffung von Software bzw. Hardware wird durch eine Checkliste erleichtert:

- a) Ist die Software gegeben, sollten für die Auswahl der Hardware folgende Fragen beantwortet werden:
 - Erstellen die Programme farbige Bilder?
 - Welche Treiber werden von den Graphikprogrammen zur Verfügung gestellt? (Werden Treiber auch von den Herstellern der Graphikkarten angeboten?)
 - Welche Graphikkartentypen unterstützen die Programme? Ist eine Graphikkarte erforderlich, die mehrere Modi unterstützt?
 - Die Fähigkeit der Graphikkarte bestimmt die Bildschirmauswahl. Entweder wird ein Bildschirm für die entsprechende Graphikkarte gewählt oder man wählt einen Bildschirm, der die verschiedenen Graphikmodi darstellen kann.

- b) Wenn die Hardware vorgegeben ist, d.h. ein PC mit Graphikkarte und Bildschirm, sind nur Programme ablauffähig, die einen Treiber für die installierte Graphikkarte haben.
- c) Wenn die Graphikhardware nachgerüstet werden soll, muß auf folgendes geachtet werden:
- Ist eine "bessere" Graphikkarte geplant, so muß auch der Bildschirm in der Lage sein, die Signale der Graphikkarte richtig darzustellen.
 - Die Graphiksoftware muß für den entsprechenden Graphikkartentyp Treiber zur Verfügung stellen.
 - Es ist teilweise bei der PC-Anschaffung schon darauf zu achten, ob die Graphikkarte integraler Bestandteil des Gerätes ist, da bei einigen neueren PCs die Graphikkarte nicht mehr ausgetauscht werden kann.

Um eine Graphik zu erzeugen, nehmen wir als einfaches Beispiel ein zweidimensionales Histogramm, muß ein Programm zur Verfügung stehen, dem die Datenwerte, die graphisch dargestellt werden sollen, eingegeben werden können. Das Programm verarbeitet die Daten in der Form, daß im zweidimensionalen Raum die notwendigen Balken in Richtung der X-Achse mit einer bestimmten Ausdehnung in Richtung der Y-Achse in sichtbarer Form dargestellt werden können. Hierzu werden etliche Rechenoperationen durchgeführt. Neben den reinen Daten "lebt" ein solches Schaubild in der Regel von einer guten Beschriftung. An den Achsen sollten die einzelnen Balken mit Labels versehen werden können. Eine Legende oder eine Bildüberschrift ist ebenfalls von Nutzen. Die Informationen für die Bildbeschriftung müssen vom Programm mit den Daten verknüpft werden. Die Bilddarstellung wirkt durch die verwendeten Formen und ihre Größenverhältnisse, die benutzten Farben und Schrifttypen. All dies ist in den Graphikprogrammen durch programmspezifische Voreinstellungen geregelt, kann jedoch beeinflussen und geändert werden. Wie leicht das geht und welche Hilfsmittel dazu genutzt werden können, macht die Benutzerfreundlichkeit eines Programms aus.

Nach den internen Rechenoperationen erzeugt das Programm eine Ausgabe. Die Programme sind zwar für bestimmte Hardware und Betriebssysteme geschrieben (z.B. IBM mit DOS), müssen aber mit den verschiedensten Hardwarekonstellationen für die Ausgabe zurecht kommen.

2.1 Geräte-Treiber

Um die Ausgabe auf einen Bildschirm, einen Drucker, einen Plotter oder ein sonstiges Ausgabemedium schreiben zu können, wurden folgende Konzeptionen entwickelt. Jede Graphik wird auf kleinste graphische Einheiten, sogenannte Primitive, reduziert. Graphische Primitive dienen zum eigentlichen Zeichnen der Bilder. Solche Funktionen sind z.B.: das Setzen der aktuellen Position auf einen bestimmten Punkt; das Zeichnen einer Linie zwischen zwei

Punkten oder eines Polygons zwischen mehreren Punkten; das Zeichnen von Texten oder das Zeichnen eines Kreises oder Kreisbogens. Programmintern werden Graphikinformationen zunächst unabhängig von der speziellen Ausgabeeinheit aufbereitet. Erst mit der Festlegung des Ausgabemediums werden diese internen Informationen für die Steuerung der entsprechenden Ausgabeeinheit umgesetzt. Diese Umsetzung erfolgt mit sogenannten Treibern. Dies sind spezielle Unterprogramme der Graphiksoftware, und sie bewirken die Umwandlung der programminternen graphischen Informationen in spezifische Steuerungsbefehle für die jeweilige Ausgabeeinheit. Da internationale Normungen fehlen, müssen jedoch je nach Herstellerfirma und Ausgabegerätetyp unterschiedliche Ansteuerungen gemacht werden, d.h. für jedes "Gerät" bzw. jeden "Gerätetyp" spezielle Treiber benutzt werden.²⁾

2.2 Graphikausgabe auf einen Bildschirm

Ein Teilgebiet der Ausgabesteuerung ist die Ausgabe auf einen Bildschirm. Die Bildschirmausgabe erfolgt nicht direkt über einen Bildschirmtreiber, sondern über eine Graphikkarte. Die heutigen PC-Bildschirme sind Rasterbildschirme, d.h. der Bildschirm ist in der Vertikalen und Horizontalen in Bildpunkte (genannt Pixel) aufgeteilt. Eine Linie am Bildschirm wird mit Hilfe einer Reihe von Punkten dargestellt. Je größer die Anzahl der Punkte in der Horizontalen und Vertikalen, desto höher ist die Bildschirmauflösung, d.h. desto weniger vermag das menschliche Auge zu unterscheiden, daß die wahrgenommene Linie nur eine Ansammlung von Punkten ist. Um ein Bild am Bildschirm in gleichbleibender Helligkeit darzustellen, muß es immer wieder neu "aufgebaut" werden. Dies muß in einer solchen Geschwindigkeit erfolgen, daß das Auge nicht den Eindruck bekommt, daß das Bild flimmert. Und schließlich spielt eine wesentliche Rolle, ob der Bildschirm Farben darstellen kann.

2.2.1 Klassifikation von Graphikbildschirmen

Der Einfachheit halber beschränken wir uns auf einige wesentliche Merkmale. Bildschirme können unterschieden werden nach:

- Farbfähigkeit; bietet der Bildschirm nur schwarz/weiß (monochrome) oder auch eine farbige Bilddarstellung?
- Bildschirmgröße, d.h. die Größe des Bildschirmdurchmessers (z.B. 12 Zoll, 14 Zoll oder noch größer).
- Auflösungsfähigkeit; sie gibt an, wie groß die Matrix der Pixel ist. Je geringer die Auflösung, desto "treppenförmiger" werden z.B. Kreise dargestellt.
- Bildfrequenz; sie ist ein Merkmal für die Flimmerfreiheit der Bilddarstellung am Bildschirm. Je höher die Bildfrequenz, desto flimmerfreier ist das Bild.
- Ausführbare Graphikmodi; es gibt Bildschirme, die nur einen einzigen Graphikmodus (z.B. EGA) korrekt verarbeiten können und solche, die mehrere Modi darstellen können und sich selbständig auf den gerade erforderlichen Modus einstellen.

2.2.2 Graphik-Karten

Die Vermittlungseinheit zwischen dem Programm und dem Bildschirm ist die Graphikkarte, d.h. eine Platine mit bestimmten Fähigkeiten, auf der sich ein Bildschirmwiederholtspeicher und eine Ausleseeinheit (Koordinierungseinheit) befinden. Für die Graphikkarten gibt es ebenfalls keine Standards; für Farbdarstellungen haben sich jedoch die IBM-Standards als Quasi-Standards herausgebildet (CGA: Computer Graphics Adapter, EGA: Enhanced Graphics Adapter, VGA: Video Graphics Adapter) und Hercules für schwarz/weiß-Darstellungen. Der Begriff Quasi-Standard meint, daß ein Kartentyp weite Verbreitung gefunden hat und von den Softwareherstellern in der Regel zu diesen Karten auch Treiber angeboten werden.

Graphikkarten lassen sich danach unterscheiden, ob und wieviele Farben sie darstellen können und welche Auflösung sie haben, d.h. die Anzahl der Bildpunkte für die Helligkeits- und Farbinformationen, die gespeichert werden können. Die Einteilung erfolgt nach der Auflösung und Anzahl der dargestellten und darstellbaren Farben.

Tabelle 1: Graphikkartentypen mit Pixelauflösung und Farbdarstellungsmöglichkeiten

Kartentyp (Firma)	Pixel-Auflösung (in vertikaler und horizontaler Richtung)	Farbdarstellung (z.B. 2 Farben aus einer Palette von 16 möglichen)
MGA (Hercules)	720x384	2/ 2
CGA (IBM)	640x200	2/16
	320x200	4/16
EGA (IBM)	640x350	16/64
VGA (IBM)	640x480	262144

Darüber hinaus gibt es etliche weitere Graphikkarten, die entweder einen dieser Standards optimieren oder mehrere Standards gleichzeitig darstellen können.

Die vom Graphikprogramm für die Ausgabe am Bildschirm über die Graphikkarte umgesetzten Informationen werden Punkt für Punkt im Bildschirm-(wiederhol)speicher auf der Graphikkarte in der Form gespeichert, wie dieser Bildpunkt auf dem Bildschirm erscheinen soll, z.B. "hell", "dunkel" und in welcher Farbe. Die Ausleseeinheit der Graphikkarte koordiniert den Zugriff auf den Bildschirmspeicher: Er muß 50 bis 70 mal in der Sekunde auf den Bildschirm ausgelesen werden, um ein flimmerfreies Bild zu erhalten, und gleichzeitig muß der Treiber sich ändernde Bilder in den Speicher einschreiben können. Je höher die Auflösung am Bildschirm sein soll, desto mehr Bildpunktinformationen muß die Graphikkarte speichern und verarbeiten können.

3. Unterschiedliche Graphik-Programmkonzepte

Schon anhand der ausgewählten Graphikprogramme ist die Konzeptionsvielfalt zu erkennen. Ein erstes Unterscheidungskriterium ist, wie vorhandene bzw. berechnete Daten für die graphische Analyse bzw. Aufbereitung zur Verfügung stehen. SIR/PC (Version 2.2.18) und SPSS/PC+ (International Version 3.0) benutzen für die graphische Darstellung eigenständige Programme. Mit SIR/PC kann eine Ausgabedatei geschrieben werden. Diese enthält die Daten und Labels, die in SIR/GRAPH wieder eingelesen und sodann als Präsentationsgraphik dargestellt werden können. D.h. um eine Graphik zu erhalten, muß SIR/PC verlassen und SIR/GRAPH (Version 4.0) als eigenständiges Programm aufgerufen werden.³⁾

SPSS/PC+ benutzt als Programme für die Graphikpräsentation über die Schnittstelle SPSS/PC+ Graphics die Produkte Microsoft CHART, Chart Master oder GrafTalk. Eines dieser Programme muß zusätzlich zu SPSS/PC+ gekauft werden. Wir beschränken uns im folgenden auf CHART (Version 3.0). Der Übergang von der Analyse in die Graphik ist über Kommandos der graphischen Schnittstelle gelöst. Das heißt, daß das SPSS/PC+ Zusatzmodul GRAPHICS benötigt wird. Hiermit ist es möglich, mit Hilfe von SPSS/PC+ Kommandos Analysewerte direkt in das Programm CHART zu übergeben. Hierzu werden die Variablen mit den Datenwerten und mit Labels in eine temporäre Hilfsdatei geschrieben, CHART von SPSS/PC+ aus aufgerufen, wobei SPSS/PC+ geladen bleibt, und die Daten der temporären Hilfsdatei werden in CHART wieder eingelesen.

SAS/PC (Version 6.03) bietet mit dem zusätzlich zu erwerbenden Modul SAS/GRAPH einen integrierten SAS-Programmbestandteil. Es enthält denselben Programmumfang, der auch bei SAS für Mainframes zur Verfügung steht. Die Graphikdarstellungen bzw. -analysen müssen nach den DATA-Steps mit Prozedurkommandos erstellt werden.

STATA/STATISTICS/GRAPHICS/DATA MANAGEMENT (Version 2.0) und STATGRAPHICS (Version 3.0) sind eigenständige Analyseprogramme. Die graphischen Darstellungen sind integrale Bestandteile der Programme. Im Programm STATA ist die Graphik mittlerweile voll integriert. Die Erzeugung von Graphiken erfolgt in STATA im wesentlichen über ein einziges GRAPH-Kommando, das viele Optionsmöglichkeiten bietet. STATGRAPHICS hingegen arbeitet menügesteuert (mit der Möglichkeit, auch Kurzkommandos einzugeben) und bietet für die verschiedensten ausgewählten Analysemethoden graphische Darstellungen.

4. Die graphischen Möglichkeiten der Programme

4.1 SPSS/PC+ Graphics mit Chart

Mit dem SPSS/PC+ Graphics-Kommando GRAPH können fünf Grundformen von farbigen Schaubildern erzeugt werden: Scatterplots (bivariate und bivariate mit einer Kontrollvariablen), Histogramme, Kuchendiagramme sowie einfache und gruppierte Balken- und Liniendiagramme. Bei den Balken-, Linien- und Kuchendiagrammen können die Werte mit Hilfe von Funktionen (z.B. kumulative Fallzahl, arithmetisches Mittel) an das Graphikprogramm übergeben werden. Im PC-Graphik-Programm CHART können dann alle CHART-Funktionen weiter eingesetzt werden. Die Hauptfunktion der Graphikanbindung an SPSS/PC+ besteht darin, gut Präsentationsgraphiken erstellen zu können. Durch die Beschränkung auf die fünf graphischen Grundformen ist die Handhabung leicht erlernbar. Um allerdings das Programm Chart in seinem vollen Funktionsumfang zu beherrschen, benötigt man weit mehr Einarbeitungszeit und Erfahrung. Als eine Schwäche hat sich die Übergabe von Daten mit größerer Fallzahl beim Einlesen in Chart erwiesen.⁴⁾ Ein Beispiel für die beschriebene graphische Schnittstelle und Chart zeigt die Abbildung 1.^{5,6)}

Die Ausgabe des SPSS/PC+ Moduls TREND II für die Zeitreihenanalyse arbeitet ebenfalls über die graphische Schnittstelle. Schließlich ist mit einem weiteren Schnittstellenprogramm SPSS/PC+ Mapping die Übergabe von Analysewerten in das kartographische PC-Programm MAP-MASTER möglich.

Hardwarevoraussetzungen: IBM PC/XT/AT, PS/2 und Kompatible, mindestens 512 K Hauptspeicher (640 K empfohlen), mindestens 10 MB Festplatte (20 MB empfohlen), unterstützt werden CGA, Hercules, EGA und VGA-Graphikkarten.

Softwarevoraussetzungen: DOS ab Version 2.0 und höher, SPSS/PC+ International Version 3.0 Base Modul.

Preise: SPSS/PC+ Graphics V 3.0 (enthält Microsoft Chart und das Graphics Interface) 1100 DM; das SPSS/PC+ Graphics Interface V 3.0 alleine kostet 600 DM.

4.2 SIR/GRAPH

SIR/GRAPH ist ein eigenständiges Präsentationsgraphikprogramm, das seine Stärken in der Darstellungsmöglichkeit vor allem dreidimensionaler Plots besitzt. Mit SIR/GRAPH können 16 vorgegebene zweidimensionale und 33 vorgegebene dreidimensionale farbige Schaubildtypen erstellt werden. Bei dreidimensionalen Schaubildern können durch die Angabe einer zusätzlichen Kontrollvariablen auch "stacked" dreidimensionale Plots, also solche mit einer vierten Dimension, gezeichnet werden. Bei dreidimensionalen Plots ist auch die Variation des Blickwinkels auf die Graphik möglich. Abbildung 2 zeigt ein Beispiel mit SIR/GRAPH.

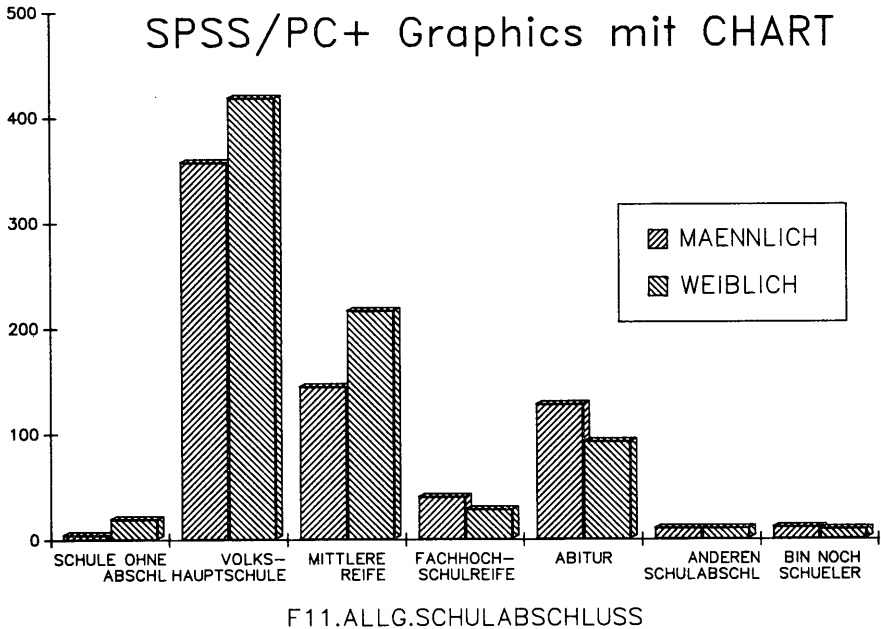


Abbildung 1

In Abbildung 1 wird die absolute Zellenbesetzung der Variablen Schulbildung (V26 mit dem Variablen-Label F11.ALLG.SCHULABSCHLUSS), aufgesplittet nach Geschlecht, in einem horizontalen Balkendiagramm mit 3-D-Effekt dargestellt.⁷⁾

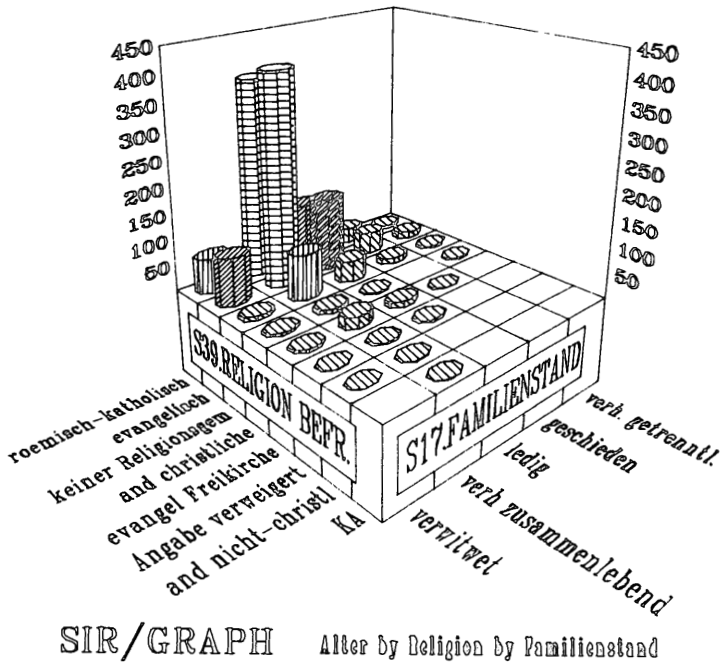


Abbildung 2

In Abbildung 2 werden die absoluten Zellenbesetzungen der Variablen Alter (V319 Alter des Befragten/kat.), aufgesplittet nach Religionszugehörigkeit (V316) und Familienstand (V237) des Befragten, in einem dreidimensionalen Balkendiagramm dargestellt, wobei die einzelnen Balken nach Häufigkeitsgruppen unterschiedliche Schraffuren haben.⁶⁾

Das Aussehen der Graphik ist mit einer Vielzahl von Manipulationsmöglichkeiten beeinflussbar. Mit Hilfe von Funktionstasten kann man zwischen einzelnen Optionen und Menüebenen wechseln. Jedoch ist die Vielzahl der Untermenüverschachtelungen ohne ausreichende Erfahrung und Einarbeitungszeit sehr verwirrend, denn auch im Manual ist an keiner Stelle die Tiefe der Menüschachtelung in übersichtlicher Form dargestellt.

Die Daten- und Labelübergabe aus SIR/PC erfolgt mit dem GRAPH SAVE FILE oder CREATE GRAPH FILE Kommando, wobei für die Aggregierung unterschiedliche mathematische Funktionen wählbar sind. Die Auswahl eines zwei- bzw. dreidimensionalen Schaubildes in SIR/GRAPH erfolgt über die Optionen ROW, COL und PAGE im GRAPH FILE Kommando. Dadurch wird eine Plattendatei erstellt. Sodann muß das Datenbanksystem SIR/DBMS oder SIR/SQL verlassen, das eigenständige Programm SIR/GRAPH gestartet und mit dem File Manager die Plattendatei geladen werden. Erst dann kann eine Graphik auf dem Bildschirm dargestellt werden.

Hardwarevoraussetzungen: IBM/PC oder Kompatibler mit DOS 2.0 und höher; mindestens 512 K RAM und eine Festplatte; unterstützt wird CGA, Hercules, EGA (mit 256 K memory), VGA und "Super EGA" (640 x 480 bzw. 752 x 410 Auflösung).

Preis: SIR/GRAPH kann nur in Lizenz benutzt werden. Die Grundlizenz für eine Kopie kostet 1 100 DM plus 176 DM für jährliche Wartung.

4.3 SAS/GRAPH

Mit SAS/GRAPH sind unterschiedliche farbige Schaubildtypen darstellbar. Im folgenden sind die verschiedenen graphischen Prozeduren kurz benannt:

- GSLIDE erstellt Textzeilen.
- GCHART erstellt vertikale und horizontale Histogramme, Kuchen-, Stern- und Blockdiagramme.
- GPLOT erstellt zweidimensionale Scatterplots mit einer Reihe von Optionen, wie die Punkte im Schaubild dargestellt werden (z.B. mit welchen Symbolen, ob und wie mit Linien verbunden oder ob eine Regressionsgerade gezeichnet werden soll).
- GMAP erstellt Landkarten in zweidimensionaler (Choropleth) und dreidimensionaler Darstellung in den Formen Surface, Block und Prism. In der Form "Surface" werden die Variablenwerte als "Gebirge" in den Landkartenregionen dargestellt. In der Form "Block" geschieht diese Darstellung mit dreidimensionalen Balken. In der Form "Prism" werden die Variablenwerte durch die unterschiedliche Höhe der Gebietseinheiten gekennzeichnet. Geographische Koordinatendateien werden bei SAS/GRAPH mitgeliefert. Für die Bundesrepublik werden beispielsweise die Landkartendaten bis auf Kreisebene zur Verfügung gestellt. Aber auch die Eingabe und Erzeugung einer eigenen Kartenkoordinatendatei ist möglich.

- GCONTOUR erstellt Contour-Plots, d.h. die Ausprägungen von drei Variablen werden in einer Fläche dargestellt.
- G3D erstellt dreidimensionale Oberflächenstrukturen (Surfaceplots) und Scatterplots. Abbildung 3 zeigt einen solchen Surface-Plot.
- G3GRID erstellt Gitternetze aufgrund von Dateninterpolationen.

Mit Hilfe einer ANNOTATE-Funktion können eigenständige Schaubilder erstellt werden.

SAS/GRAPH ist kommandogesteuert. Benutzt man die vordefinierten Standardwerte, ist recht schnell ein Schaubild gezeichnet. Will man jedoch darüber hinaus die Voreinstellungen manipulieren, muß wegen der komplexen Möglichkeiten eine längere Einarbeitungszeit berücksichtigt werden. Als negative Eigenschaften schlagen zu Buche, daß bei größeren Datenmengen und Schaubildmanipulationen die 640 K RAM schnell erschöpft sind, so daß auch vom SAS Institute empfohlen wird, den PC mit einer expanded memory Karte mit zusätzlichen 2 MB auszurüsten.⁹⁾ Schließlich benötigt SAS/GRAPH im Vergleich zu Statgraphics oder STATA auch mehr Rechenzeit für die Bildaufbereitung.¹⁰⁾

Hardwarevoraussetzungen: IBM PC oder Kompatibler mit 640 K RAM; unterstützt werden CGA und EGA.

Softwarevoraussetzungen: DOS 2.0 und höher sowie das SAS BASE SYSTEM Version 6.03.

Preise: SAS/GRAPH kann nur in Lizenz genutzt werden. Eine Einzelkopie kostet im Anschaffungsjahr 1200 DM und in den Folgejahren pro Jahr 600 DM.

4.4 STATA/GRAPHICS

STATA ist ein sehr kompaktes Statistikprogrammpaket mit integrierter Graphik. Die statistischen Berechnungen und Graphikaufbereitungen sind auch bei größeren Datenmengen sehr schnell. Es ist für Statistiker und Sozialwissenschaftler angelegt, die die Graphik nicht nur als reine Datenpräsentation, sondern auch zur statistischen Datenanalyse einsetzen wollen.

STATA/GRAPHICS stellt acht farbige Grundtypen von Graphiken zur Verfügung: Histogramm, eindimensionaler und zweidimensionaler Scatterplot, Scatterplot Matrix, Box-and-Whisker-Plot, Star-Plot, Balken- und Kuchendiagramm. Diese Graphiken können auch untereinander kombiniert in einem Schaubild dargestellt werden, z.B. ein eindimensionaler Scatterplot verbunden mit einem zweidimensionalen Scatterplot und einem Box-and-Whisker-Plot. In Abbildung 4 sind zweidimensionale Scatterplots mit Box-and-Whisker-Plots kombiniert. Die Box in der Mitte beschreibt die Verteilung der mittleren 50% der Datenwerte mit dem Median (mittlerer Strich) sowie dem unteren und oberen Quartil, gekennzeichnet durch die untere und obere Boxbegrenzung. Die

SAS/GRAPH

dreidimensionaler Surfaceplot
Alter x Schulbildung x Nettoeinkommen

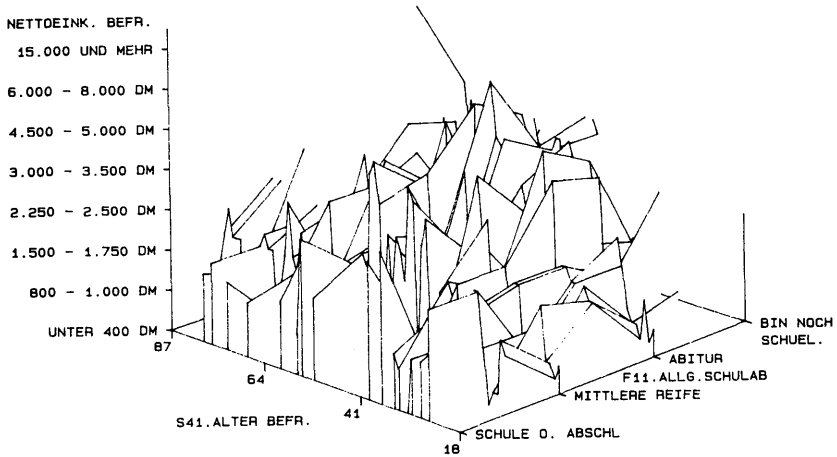
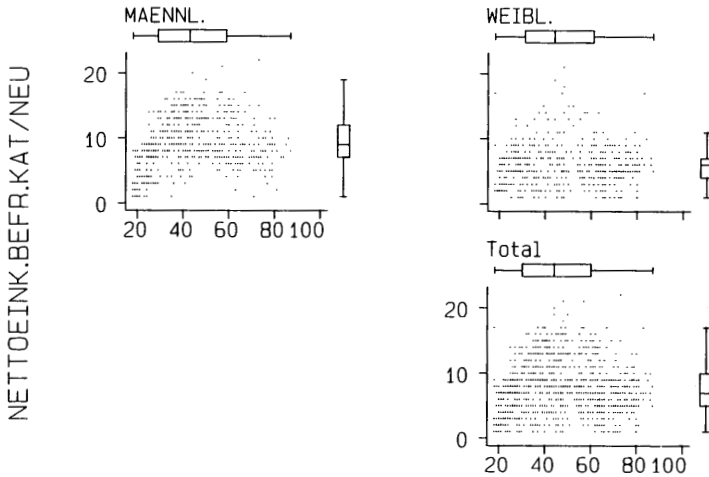


Abbildung 3

In Abbildung 3 werden die Punkteverteilungen der Variablen Alter (V318 Alter des Befragten), allgemeiner Schulabschluß (V26) und Nettoeinkommen (N373 kategorisiert) im dreidimensionalen Raum gezeichnet und diese Punkte untereinander verbunden, so daß eine Oberflächenstruktur dargestellt wird, die einen Eindruck über die Tendenz des Zusammenhangs von Alter, Schulbildung und Nettoeinkommen erkennen läßt.¹¹⁾



S41.ALTER BEFRAGTER
STATA – Graphs by S1.GESCHLECHT BEFR.

Abbildung 4

In Abbildung 4 werden die Werteverteilungen der Variablen Alter des Befragten (V318) und Nettoeinkommen kategorisiert (N373) in ihrer Gesamtdarstellung (TOTAL) sowie aufgesplittet nach Geschlecht (V196) in zweidimensionalen Scatterplots dargestellt. Als zusätzliche statistische Information über die Punktwolken hinaus wurden die Box-and-Whisker-Plots für die Variablen Nettoeinkommen und Alter pro Scatterplot hinzugezeichnet.¹³⁾

von der Box ausgehenden Striche und deren Länge bezeichnen die Datenwerte für das Minimum und Maximum.

Aus der Fülle der möglichen Darstellungsvarianten sollen im folgenden nur einige hervorgehoben werden: Histogramme können automatisch mit einer Normalverteilung überlagert werden. In einem Scatterplot können mehrere, unterschiedliche Variablen abgebildet werden. Eine Stärke von STATA sind die Matrix-Scatterplots: Bis zu 9*9 Variablen können gegeneinander in einem Plot dargestellt werden. Auch die gewichtete Darstellung von Datenwerten in einem Scatterplot ist möglich. Mittels eines Zusatzmoduls GRAPH.KIT - eine Sammlung von zusätzlichen STATA-Programmen, die mit STATA-Kommandos erstellt wurden - können weitere sieben spezielle Plots aufgerufen werden (z.B. error-bar-chart, highlight a subset of points in an two-way-graph). Eine Schwäche von STATA/GRAPHICS sind die geringen Möglichkeiten der Präsentation¹²⁾aufbereitung.

Hardwarevoraussetzungen: IBM PC oder Kompatible mit mindestens 256 K RAM mit DOS 2.0 und höher, zwei Diskettenlaufwerke oder eine Festplatte; unterstützt wird expanded memory (EMM) sowie die Graphikkartentypen Hercules, CGA, EGA und VGA.

Preis: STATA kostet 665 Dollar plus 30 Dollar Transportkosten, wenn man es von Europa aus direkt in den USA bezieht.

4.5 STATGRAPHICS

Das Programm STATGRAPHICS ist so aufgebaut, daß für eine sehr große Menge statistischer Verfahren die Darstellung farbiger Graphiken möglich ist. Hierin liegt die entscheidende Stärke des Programms. Es ist menügesteuert und dadurch relativ leicht handhabbar und erlernbar. Man kann aber auch mit Kurzkommandos über Menüebenen hinwegspringen. Ähnlich wie STATA ist STATGRAPHICS nicht als reines Präsentationsgraphikprogramm konzipiert. Allerdings sind im Gegensatz zu STATA mehr Einflußmöglichkeiten auf Präsentationseffekte gegeben. In Abbildung 5 wird ein "Hanging Histogram" dargestellt.

Hier können nicht alle Möglichkeiten beschrieben werden. Im folgenden werden nur die Oberbegriffe der verschiedenen Verfahrensmöglichkeiten benannt und am Beispiel der Verfahren für "ANOVA and Regression Analysis" die Darstellungsformen im einzelnen aufgelistet. An Verfahren stehen zur Verfügung:

- PLOTTING AND DESCRIPTIVE STATISTICS mit: Plotting Functions/Descriptive Methods/Estimation and Testing/Distribution Functions/Exploratory Data Analysis
- ANOVA AND REGRESSION ANALYSIS mit Analysis of Variance und darunter:
 - . One-Way Analysis of Variance
 - . Multifactor ANOVA

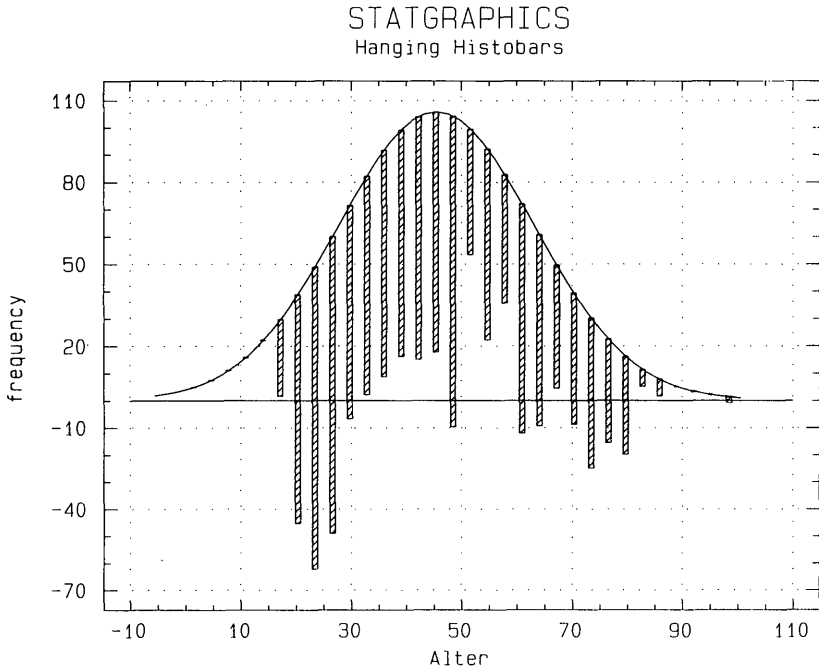


Abbildung 5

Abbildung 5 zeigt "Hanging Histograms" der Variablen Alter des Befragten (V318).¹⁵⁾ Es ist ähnlich wie ein Histogramm aufgebaut, nur werden die Häufigkeiten der einzelnen Zellenbesetzungen ausgehend von der am besten angepaßten Normalverteilung nach unten abgetragen. Sind auch die Häufigkeiten der Datenwerte normalverteilt, so schwanken die Enden des Histogramms geringfügig um die Nullpunktlinie. Der Grad der Abweichung der Datenwerte von den erwarteten Häufigkeiten wird deutlich, wenn die Nullpunktlinie nicht erreicht oder überschritten wird.

- . Analysis of Nested Designs
- . Kruskal-Wallis One-Way Analysis by Ranks
- . Friedman Two-Way Analysis by Ranks
- sowie Regression Analysis und darunter:
 - . Simple Regression
 - . Interactive Outlier Rejection
 - . Multiple Regression
 - . Stepwise Variable Selection
 - . Redge Regression
 - . Nonlinear Regression
- TIME SERIES PROCEDURES mit: Forecasting/Quality Control/Smoothing/Time Series Analysis
- ADVANCED PROCEDURES mit: Categorical Data Analysis/Multivariate Methods/Nonparametric Methods/Sampling/Experimental Design

Zwar ist Statgraphics bei Berechnungen schnell, ob überhaupt ein Ergebnis berechnet werden kann hängt jedoch bei größeren Datenmengen entscheidend davon ab, wieviel Kernspeicher für die Berechnungen noch zur Verfügung steht und wieviele Variablen mit welcher Fallzahl in die Analyse einbezogen werden. Je nach Verfahren, bei denen eventuelle temporäre Variablen benötigt werden, ist dann schnell der zur Verfügung stehende Kernspeicherbereich erschöpft und die Meldung "insufficient RAM memory" bricht die Berechnungen ab.¹⁴⁾

5. Zusammenfassung

Wie aus den Beschreibungen der verschiedenen Graphikkonzepte und den Beispielprogrammen ersichtlich wurde, kann eine Empfehlung für ein Produkt ohne die Beachtung der Rahmenbedingungen, in die die Graphik eingebunden werden soll, nicht gegeben werden. Je nach Anwendungsgebiet, ob mehr die Präsentationseigenschaft oder mehr die interaktiven Möglichkeiten mit Graphik genutzt werden sollen, wird eine Entscheidung anders ausfallen müssen und werden die eventuellen Nachteile für das gewählte Programm daraufhin abgewogen werden müssen, ob sie für den praktischen Einsatz noch tragbar sind. Für die Zukunft sind ständige Weiterentwicklungen angekündigt. Hierzu zählen z.B. die Integration von Fenstertechniken, der Einsatz einer Maus oder auch die Einbindung von Graphik in Text. Es wird sich weiterhin lohnen, den Markt aufmerksam zu beobachten.

Anmerkungen

- 1) Auch bei den Dateneingabemöglichkeiten ist man vielfach nicht mehr auf die klassische Tastatur beschränkt. Die Anschlußmöglichkeiten einer Maus, eines Graphiktableaus bis hin zu Scannern eröffnen auch hier neue Möglichkeiten.
- 2) So geschieht z.B. die Umsetzung einer graphischen Ausgabe für einen HP-Plotter in die HP-spezifischen Kommandos HPGL (Hewlett Packard Graphic Language), die der Plotter versteht und in Stift- und Papierbewegungen umsetzt.

- 3) Bei SIR/GRAPH handelt es sich um eine Weiterentwicklung des Programms BoeingGraph.
- 4) Im konkreten Fall handelte es sich um eine Datei mit 1500 Fällen (siehe auch Anmerkung 5). Es sollte ein Scatterplot erstellt werden mit dem Kommando "GRAPH/SCATTERPLOT N373 BY V318" (N373 = Nettoeinkommen des Befragten, V318 = Alter des Befragten). Zum Einlesen in Chart wurde mit einem IBM/AT03 mit einem 80286 Prozessor, einem mathematischen Coprocessor, 640 KB RAM und einer EGA-Karte etwas mehr als 5 Minuten benötigt, um 3075 Zeilen aus der Datei SPSS.SLK, der Zwischenspeicherdatei für die Datenpunkte der Graphik zu lesen. Danach hängte sich das System auf, und nur mit einem Warmstart konnte Chart wieder verlassen werden. Dasselbe Phänomen ereignete sich mit einem IBM PS/2 Modell 50, allerdings wurden die Datenwerte hier um ein Drittel schneller eingelesen. Der Scatterplot konnte mit Chart nur auf dem Bildschirm dargestellt werden, wenn das Programm Chart alleine, also nicht über SPSS/PC+, geladen war und die Datei SPSS.SLK eingelesen wurde.
- 5) Datengrundlage ist eine Stichprobe mit 1500 Fällen und 11 Variablen aus dem ALLBUS (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften) von 1986.
- 6) Aus drucktechnischen Gründen konnten die Schaubilder nur in schwarz-weiß reproduziert werden. Sie wurden mit einem Tischplotter HP 7550A gezeichnet.
- 7) Das SPSS/PC+ Kommando lautet: "GRAPH/BAR = COUNT BY V26 BY V196". Die Schaubildbeschriftung wurde in Chart hinzugefügt.
- 8) Das SIR/PC-Kommando für die Erstellung der Ausgabeplattendatei für SIR/GRAPH lautet: "GRAPH SAVE FILE OBS = (COUNT V319) / ROW = V316 / COL = V237"
- 9) So ist es uns bei einem IBM/AT03 mit 640 KB RAM ohne expanded memory Karte nicht möglich, die Landkarten-Kreisdaten der BRD überhaupt vollständig einzulesen oder das abgedruckte Beispielschaubild kann nicht mehr ausgegeben werden, wenn bei den GOPTIONS HTEXT=50 PCT gesetzt werden soll, um die Schriftgröße zu beeinflussen.
- 10) Im aktuellen Beispiel benötigte die Prozedur G3D drei Minuten für die internen Berechnungen.
- 11) Die SAS-Kommandos für dieses Schaubild lauten:
 GOPTIONS DEVICE=HP7550A HSIZE=7 IN VSIZE=5 IN COLORS=(BLACK);
 TITLE1 F=SWISS 'SAS/GRAPH';
 TITLE2 F=COMPLEX 'dreidimensionaler Surfaceplot';
 TITLE3 F=COMPLEX 'Alter x Schulbildung x Nettoeinkommen';
 PROC G3D DATA=ALLBUS86;
 FORMAT V26 V26FMT.;
 FORMAT N373 N373FMT.;
 PLOT V318 * V26 = N373 / ROTATE=45 ZTICKNUM=8 SIDE;
 RUN;
 Mit SAS/ASSIST können auch menügesteuerte Graphiken erstellt werden.
- 12) So können z.B. Überschriften nicht an beliebige Stellen im Schaubild verschoben werden. Im Beispielschaubild konnten die x- und y-Achsen auch nicht mit Labels versehen werden. Es ist jedoch möglich, die STATA/GRAPHICS-Ausgabeplattendatei in das Lotus Freelance Präsentationsgraphikprogramm einzulesen und dort weiter aufzubereiten.
- 13) Die STATA-Kommandos für dieses Schaubild lauten:
 SET TEXTSIZE 150
 GRAPH N373 V318, BY(V196) TOTAL TWOWAY BOX YLAB XLAB SYMBOL(.)
 TITLE("STATA - Graphs by S1.GESCHLECHT BEFR.")
- 14) Mit unseren Beispieldaten (1500 Fälle) konnte die fiktive Berechnung einer einfachen Regression mit der abhängigen Variablen V318 und der unabhängigen Variablen V196 nicht durchgeführt werden.
 Im Statgraphics Manual werden die folgenden Hilfsstrategien für solche Probleme angeboten:
 - "Increase the Memory Threshold in your System Profile";
 - "Use Lotus Intel Memory to hold approximately 50K of Statgraphics functions";

- "Use a RAM disk or hard disk for swapping out currently unused Statgraphics routines and data using the Extended Work Area".

- 15) Das Schaubild wurde erstellt durch die Statgraphics Menüauswahl "G. Estimation and Testing" und sodann "4. Hanging Histograms". Im "Tabulation Input Panel" wurden die Schaubildüberschriften und die Achsenbeschriftungen eingetragen.

Literatur

Computer Resource Center, 1988: STATA Reference Manual. Los Angeles.

Microsoft Corporation, 1987: Microsoft Chart. Redmond.

Purgathofer, Werner, 1986: Graphische Datenverarbeitung. Wien, New York (2., verb. Aufl.).

SAS Institute Inc., 1987: SAS/GRAPH Guide for Personal Computers, Version 6 Edition. Cary.

SAS Institute Inc., 1988: SAS/GRAPH Hardware Interfaces for Personal Computers, Version 6 Edition. Cary.

SIR, a Division of ISI, 1988: SIR GRAPH Reference Manual. Deerfield.

SPSS Inc., 1986: SPSS/PC+ Graphics V2.0 for the IBM PC/XT/AT and PS/2. Chicago.

STSC Inc, 1986: STATGRAPHICS User's Guide, Part 1 and 2. Rockville.